# Identity Management for Interoperable PTC Systems in Bandwidth-Limited Environments

# The Final Report

## Part 1 (of three parts)

## Introduction to The Project

## And

## Performance Studies

Prepared for the Federal Railroad Administration
Under Grant Number FR-TEC-0006-11-01-00/20.321

Version prepared on September 28th, 2014 as a revision to the version from May 2014.

Principal Investigator: Prof. Rajni Goel, Howard University

Co-Principal Investigators: Prof. Duminda Wijesekera, George Mason University

Dr. Andre B. Bondi, Siemens Corporation

## Table of Contents

## List of Figures

## List of Tables

## Executive Summary

Positive Train Control is a wireless based system designed to provide comprehensive safety coverage for passenger and cargo trains operating on US railroads by 2015. Mandated by Rail Safety Improvement Act of 2008 (RISA 2008), major railroads have designed a broad architecture consisting of two networks; namely the Signaling Network (SN) and the

Wayside Interface Unit (WIU) network powered by software-defined radios (SDRs) that use the same 220MHz range. The Signaling Network provides authorities for trains to enter fixed blocks of track and other signal functions and the Wayside Interface Network provide sensory information about the vicinity of the tracks. The railroad community has decided that both network require message integrity and availability but not confidentiality for both networks.

The research work originally funded was going to address the ability to create an identity management system for PTC signaling, covering both the Signaling network and the WIU network. Subsequently, as the work progressed, the project scope concentrated on the security of the WIU network.

Our main findings show the need to propose an enhanced WIU network. The second part articulates the need and the third part describes a potential solution and a prototype implementation.

# Section 1. The Project Reports

This report is prepared as a final deliverable for the Grant Number FR-TEC-0006-11-01-00/20.321. The project report consists of three parts. This is the first part, and it describes the project objectives and evolution that produced the work reported in Parts 2 and Part 3 and performance measures done for the WIU network. Part 2 describes our understanding of the WIU network and Part 3 provides a design and prototype implementation details of the WIU protocols security.

The rest of the first part of this report is organized as follows. Section 2 describes the project objectives and how they changed during the project execution with the approval of the FRA. Section 3 summarizes the contents of Parts 2 and 3. Section 4 describes the performance studies undertaken under this grant. Section 5 describes how we estimated the bandwidth requirements between the On Board Unit (OBU) and the Back Office (BO). Section 6 has our concluding comments.

# Section 2. Project Objectives and Progress

The initial emphasis our work was on security and performance aspects identity management for trains moving between networks. During the course of our work, it has become clear that I-ETMS will be the dominant standard for Positive Train Control in most of the USA, except on Amtrak's Northeast Corridor, where it will coexist with ACSES2 when freight and passenger service run on the same tracks. I-ETMS does not support handshaking authentication of trains *per se.* Instead, the validity of the control messages exchanged between the signaling system and the trains rests on two pillars, the security of a database on board each train, and the integrity of the messages exchanged between a train and two key components of the signaling system, the Wayside Interface Units (WIUs) and the Back Office Servers (BOSs).

- Our understanding of the published documents say that s trains' requests for wayside status messages are assumed to be valid if they have valid HMAC keys because the keys are stored in a secured on-board database. The keys may pertain to the railroad on which the train is running or to the railroad the train seeks to enter.
- Trains rely on the integrity of the messages (validated using an HMAC field with pre-stored keys) to determine that the messages came from an authentic source.
- Therefore, the emphasis of our work has moved away from authentication towards the WIU messages is their integrity, (as the messages are not encrypted) and the overall security of the beaconing. This change of emphasis reflects our changing understanding of PTC and was communicated to and agreed to by the FRA in face-to-face meetings, conference calls, and e-mails.
- The list of corresponding deliverables has changed along with our evolving understanding of the research issues and with our and the FRA's changing concerns about the safety, security, and performance.

- On a subsequent meeting with Mr. David Balckmore and Dr. Mark Hartong, we demonstrated the details of the vulnerabilities and how they may be exploited. Subsequently, the FRA contacted Metorcomm to show these vulnerabilities with their radios and George Mason University and Howard University got three radios each for the train, back office and the WIU.
- We also sent one of GMU graduate students for training to the Meteorcomm facility for two classes for radio setup and usage. This activity took place during the first quarter of calendar year 2012. Our subsequent experiments with the radio sets revealed that the radios were given to us without any software to operate them. We had numerous conversations with Meteorcomm to obtain this software for experiments, but they refused saying that their radios had no known vulnerabilities and the two schools did not understand how hashing works and if anything at all needs to be tested, we should give our source code to Meteorcomm for testing. In our professional opinion, this violates all safeguards placed in securing and security testing software, hardware and devices.
- As a consequent, with the approval of the FRA, GMU-Howard team decided reproduce the WIU messaging standard and demonstrate the ability to exploit the published WIU broadcast protocol and a potential strengthening thereof, followed by an implementation and a running demonstration. For this part the GMU-Howard team provided three demonstrations that successively built upon the previous demonstrations showing the vulnerability of the existing WIU protocol. The first stage demonstrated a WIU become running on two computers but used the TCP-IP protocols as a delivery mechanism of the packets. The second stage used two plug computers that used the Wi-Fi protocol for message transmission. The third stage consisted of two Ettus radios that naively transmitted WIU messages using the BPSK encoding of WIU messages.

## Section 3. Contents of Part 2 and Part 3

Part 2 of this report provided separately describes our understanding of the proposed WIU network for operational PTC systems. This is provided as a stand-alone document.

Part 3 of the this report provided separately describes a WIU broadcast system designed and prototyped using Software Defined Radios produced by Ettus Inc.

## Section 4. Performance Studies

The performance studies reported in this section were lead by Dr. Andre Bondi of Siemens Corporate Research and assisted by Damindra Bandara, a graduate student from George Mason University. The written report on this topic has been constructed from written reports provided as a deliverable from Siemens.

### 4.1. Quantities of Interest in Performance Requirements

Generally, delay and throughput are the main variables of performance requirements.

- *Delay.* In PTC, we are interested in ensuring that train control messages are delivered early enough to affect movement and speed of trains if drivers do not react to signals or speed resrictions properly. The available time to do this depends on train speed, the lengths of blocks, the positions of signals and switches, radio range, and the braking properties of trains. The ability to meet delay requirements depends on the overall volume of train movements, the numbers of messages associated with each train movement, the numbers of interactions with the PTC system resulting from each train movement, and the resource demands of each type of interaction.
- *Throughput.* The demand on a PTC system is affected by the number of trains it must control during the peak hour and the number of switches and signals that the trains will pass during that hour. Bandwidth and the computers controlling the Wayside Interface Units (WIUs), On-Board Units (OBUs), and Back Office Servers (BOSs) all influence the PTC system's ability to meet this demand.

In the remainder of this section, we examine the relationship between braking properties and delay requirements and then go on to discuss how system throughput is driven by track layouts and timetables.

## 4.2. Braking Curves, Train Speed, and Distance Traveled

### 4.2.1. Time Available to Respond to a Command

The time available for a train to respond to a command depends on the current distance from the train to the place at which the train must respond to a command, the reaction time of the engine driver, the speed at which the train is moving, the time taken for PTC to intervene if the driver has not acted upon a signal, and the shape of the braking curve. The braking curve represents the speed of the train as a function of distance traveled when specific actions are supposed to occur, such as the application of brakes or the issuance of a stop signal or a speed restriction signal.

### 4.2.2. Moving from Braking Curve to Delay Requirement

Computer performance requirements are usually expressed in terms of throughput or offered load and response time. Locomotive braking curves are typically show speed as a function of distance rather than time [Groepler20xx, Vincze2006, Wei2010, Hunt2011].

Specification of delay requirements for delivering emergency stop signals or notification that previously blocked track is now open depends on the shape of the braking curve, the performance characteristics of the OBU, the characteristics of the communications network, and on operating procedures. The advantage of delivering notification that a section of track has opened is that the train need be slowed to a stop, only to be required to accelerate from rest. This can result in considerable energy savings, as well as reducing wear and tear on brake shoes, wheels, the tracks, and the rail bed.

### 4.2.3. Computing Deceleration Times from a Braking Curve

Figure 1 shows an example of a braking curve. *D* is the desired stopping point. *E* is the actual stopping point, which ideally should be in the same position as *D* or even to the left of it. The situation as shown in Figure 1 is undesirable. The train has clearly overshot the signal. The points shown on the X axis are as follows:

- *D* is the desired stopping point. It is some distance before the home signal.
- *E* is the actual stopping point.
- *R* is the point at which the signal is shown to be at danger inside the cab.
- *M* is the position of the distance signal mast. The distance signal indicates the aspect of the home signal.
- *B* is the point at which the engine driver starts to activate the brakes.
- *P* is the point at which PTC braking is activated if the engine driver has failed to stop the train after passing the distance signal mast *M.*

The relative positions of these points are determined by the anticipated maximum speed of the train and the shape of the braking curve. A steep braking curve, such as the service braking curve or emergency braking curve, is undesirable, as it causes brake wear, might flatten wheels on one side and/or cause track damage, and could lead to any or all of passenger discomfort, personal injury, and damaged cargo.

If a message arrives saying that the signal just after *D* is CLEAR when the train is between points *A* and *P*, Positive Train Control could be used to send a corresponding message to the cab and cancel the order to stop. This could result in shorter travel time and energy savings.

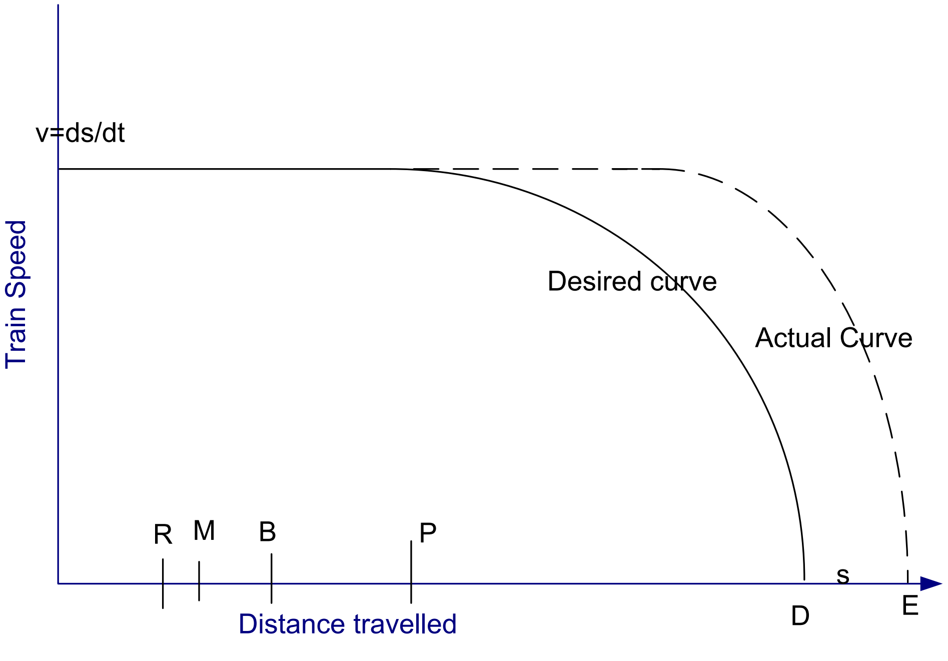

**Figure 1. An example of a braking curve.**

To overcome the difficulties of deriving a delay requirement from a curve expressed in terms of speed and distance, we first note that the braking curve can be expressed as an equation of the following form:

$$\boldsymbol{\frac{ds}{dt} = f(s)} \tag{0.1}$$

$s$ is the distance travelled, and $f(s)$ is a (sometimes proprietary) braking curve satisfying $f(0) = v_0, f(D) = 0$. If the brakes are not activated, as would be the case up to distance $A$, we would have

$$\boldsymbol{f(s) = v_0, 0 \leq s \leq B} \tag{0.2}$$

At the point where the brakes are activated, we can write

$$f(s) = b(s)$$

$b$ is continuous and non-increasing, with the desired properties that

$$b(A) = v_0\, , b(D) = 0.$$

If the driver passes a signal at danger, $ds/dt = v_0$ for some $s' \geq B$.

Without loss of generality, we can use the technique of separation of variables to write

$$\boldsymbol{\frac{ds}{f(s)} = dt} \tag{0.3}$$

From this, we obtain the stopping time $T = t_E - t_R$ by writing

$$\boldsymbol{\int_{s=R}^{s=E} \frac{ds}{f(s)} = \int_{t_R}^{t_E} dt = t_E - t_R} \tag{0.4}$$

and solving for $t_E$. Delay requirements over the various track segments *RM, MA, PD,* and *PE* are defined analogously. Thus, our problem becomes (a) the identification of these points of action and (b) the corresponding required delay times. These may be mutually dependent. They should be engineered for the maximum anticipated train speed and the minimum anticipated distance between points *R* and *D.*

The left hand side of equation (0.4) has a zero denominator at *s=E.* This is potentially problematic when evaluating equation (0.4). A related singularity problem when analyzing braking curves is described in [Hunt2011], as we shall see below.

We shall circumvent this problem by using the physical properties of the system to justify the solution of equation (0.1) by numerical means instead. Specifically,

1. *f(s)* is given either explicitly or in numerical form. We are trying to find the value of *t* (not *s*) such that *f(s)*=0.
2. *f(s)* is continuous and differentiable with respect to *s* in the range of interest. This is simply a way of stating that the braking curve is smooth and that there will not be sudden changes in speed.
3. Similarly, the speed of the train evolves smoothly with respect to time, as does its position. The train does not suddenly jump from one place to the next. Formally, if *x* is the displacement of the train, we have from the chain rule:
$$\boldsymbol{\frac{ds}{dt} = \frac{ds}{dx}\frac{dx}{dt}} \tag{0.5}$$
4. Equation (0.1) is in *autonomous* form, because the independent variable $t$ does not appear on the right hand side.

Thus all derivatives involved are continuous if the original braking curve is.

The analysis also applies to cases in which the train is allowed to proceed past a signal at restricted speed *r*. In that case, the goal of the braking curve is to bring the train speed down to $r < v_0$. Then, without loss of generality, we write

$$f(s) = r, s \geq D; r < v_0$$

$$b(A) = v_0, b(D) = r\ .$$

We now turn to the numerical solution of equation (0.1). Note that travel time is distance divided by the average speed. Marking the $s$ axis with points $s_0, s_1, \ldots, s_K$, suppose that the locomotive is at point $s_i$ at time $t_i$. Since the speed along the braking curve is non-increasing, the average speed between points $s_i$and $s_{i+1}$ may be approximated by the average of the speeds at these two points. Thus, we can write the following approximation:

$$\boldsymbol{t_{i+1} \approx t_i + \gamma\ \frac{s_{i+1} - s_i}{[f(s_{i+1}) + f(s_i)]/2}\ , i = 0, 1, \ldots, K - 1} \tag{0.6}$$

where γ is a multiplicative constant used to ensure that the units are correct. This midpoint method is an adaptation of methods found in [DahlBjork1974]. In the USA, the denominator in the second term will be expressed in miles per hour and the numerator in feet. Thus, $\gamma = 3600/5280$. In countries using the metric system, the denominator will be expressed in kilometers per hour and the numerator in meters. Thus, in the metric system, $\gamma = 3600/1000$.

Notice that as the train slows, with *f(s)* approaching zero, the times between reaching equally spaced successive positions $s_i, s_{i+1}$ will increase, as we would expect. Notice also that the denominator in equation (0.6) will never be zero because $f(s_{K-1}) > 0$ even if $f(s_K) = 0$, by construction.

The approach we have just described is applicable to generalized braking curve functions, whether they are explicitly specified or generated numerically based on circumstances such as the slope and/or curvature of the track. The braking algorithm giving rise to the curve $f(s)$ for some locomotives and control systems may not be in the public domain. Eventually, vendors may wish to use our suggested method to evaluate stopping times as well as braking distances. Ideally, when undertaking our numerical investigation of performance requirements for PTC delays, we would like to examine one or more published braking curves so that one can see the impact of braking algorithms on the PTC requirements. As we have not found any published braking curves at the time of writing, we shall begin our numerical investigation by resorting to a simple hypothetical braking curve *f(s)* with the following properties:

1. $f(s) = v_0, s \leq B$
2. $f(D) = 0$
3. $f^{'}(B^{-}) = f^{'}(B^{+}) = 0$
4. $f^{'}(s) < 0, B < s \leq D$
5. $f^{''}(s) < 0, B \leq s \leq D$

The first two conditions are initial and terminating conditions. The third condition implies that *f* is differentiable at *s=B*, meaning that there is no sudden jump in the rate at which the train speed changes with respect to distance traveled. The fourth condition states that *f* is decreasing once the brakes are applied, meaning that the train will slow down, while the fifth condition implies that *f* is twice differentiable over the specified range and concave down, meaning that braking becomes sharper with respect to distance as the train approaches the required stopping point.

We shall construct a quadratic polynomial to act as a synthetic form for *f* for $\leq s \leq D$ . Any function with properties 1-5 would do, but a quadratic is particularly simple to obtain. If the polynomial takes the form

$$f(x) = c_2x^2 + \ c_1x + c_0 \tag{0.7}$$

We require that

1. $c_2B^2 + \ c_1B + c_0 = v_0$
2. $c_2D^2 + \ c_1D + c_0 = 0$
3. $2c_2B + c_1 = 0$

We proceed by back-substituting the third condition into the first two, and then solving for $c_2$ and $c_0$ using Cramèr's rule [Cohn1958]. We have $c_1 = \ -2Bc_2$ immediately, yielding

$$-B^2c_2 + c_0 = v_0 \tag{0.8}$$

$$(D^2 - 2BD)c_2 + \ c_0 = 0 \tag{0.9}$$

Cramèr's rule for determinants followed by substitution in condition 3 for $c_1$ gives us

$$c_2 = \frac{\begin{vmatrix} v_0 & 1 \\ 0 & 1 \end{vmatrix}}{\begin{vmatrix} -B^2 & 1 \\ D^2-2BD & 1 \end{vmatrix}} = \frac{-v_0}{B^2-2BD+D^2} = \frac{-v_0}{(D-B)^2}$$

$$c_0 = \frac{\begin{vmatrix} -B^2 & v_0 \\ D^2-2BD & 0 \end{vmatrix}}{(D-B)^2} = \frac{(D^2-2BD)v_0}{(D-B)^2} \tag{0.10}$$

$$c_1 = \frac{2Bv_0}{(D-B)^2}$$

Hence, we have

$$f(x) = \frac{v_0}{(D-B)^2}[-x^2 + 2Bx + \ (D^2 - 2BD)], B \le x \le D. \tag{0.11}$$

Notice the form of the coefficients: all are inversely proportional to the square of the desired braking distance, and all are proportional to the initial speed. The quadratic coefficient and the constant term are negative. Checking that the speed will be zero at *x=D,* we have

$$-D^2 + 2BD + (D^2 - 2BD) = \ 0$$

as required. Checking that the speed will be $v_0$ at *x=B*, we have

$$\frac{-v_0B^2 + 2B^2v_0 + (D^2 - 2BD)v_0}{(D-B)^2} = v_0$$

again as required.

Following Tse and Brosseau [TseBross2009], we can then evaluate the situation with $v_0 = 40$ mph and $D - B \sim 6000$ feet. Without loss of generality, we take *B*=0 feet and *D*=6000 feet. Substituting into equation **(0.11)**, we have

$$f(x) = \frac{40}{3.6\times10^7}[-x^2 + \ 3.6\times10^7], 0 \le x \le 6000$$

We obtain the plot shown in Figure 2.

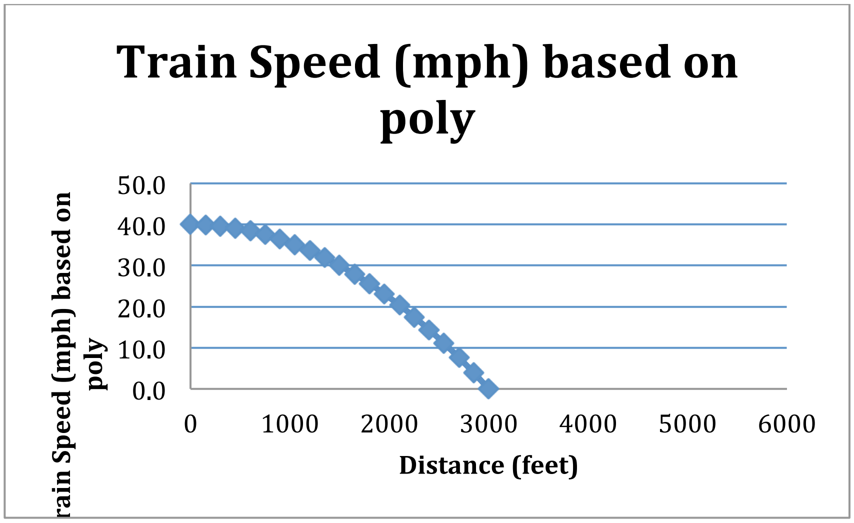

**Figure 2: A Synthetic Braking Curve.**

Note: if a sharper gradient is required near the stopping point, we can set the desired values of the first derivative there, but this will entail fitting cubic polynomial instead of a quadratic one, because we need to add more coefficients to allow more degrees of freedom as constraints are added. We can also come up with families of braking curves, such as the emergency braking curve, by specifying differing values of

Figure 3. shows the estimated time taken to travel the corresponding distances, based on equation (0.6). The curve shows that when the stopping distance is just over a mile, the train takes just under five minutes to come to a complete stop from 40 mph. A much steeper braking curve with a shorter stopping distance is needed if the train is to be used on a commuter line with stops a few minutes apart. Hence, the quadratic braking curve we have used is not acceptable for this type of train. On the other hand, stopping distances of a mile or more for very long and heavy freight trains are not unheard of. This underscores the value of delivering a timely go signal to the cab if the driver was previously slowing the train down in anticipation of a stop signal. Avoiding a complete stop saves braking effort, acceleration time, and energy, quite apart from reducing wear on the brakes. Since trains are programmed to stop by default in the absence of other information, timely delivery of CLEAR messages is particularly useful.

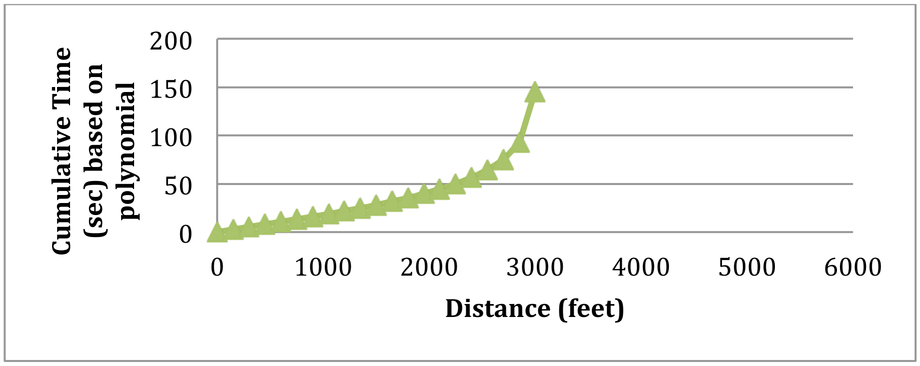

**Figure 3. Cumulative Travel Time from Brake Application at Position Zero.**

Notice that points on this graph are equally spaced, and that the time taken to travel between them is estimated to be longer in each successive interval, because the train has slowed down. The inverse is shown in the next figure. Here, we see that the train decelerates markedly as it approaches the stop signal.

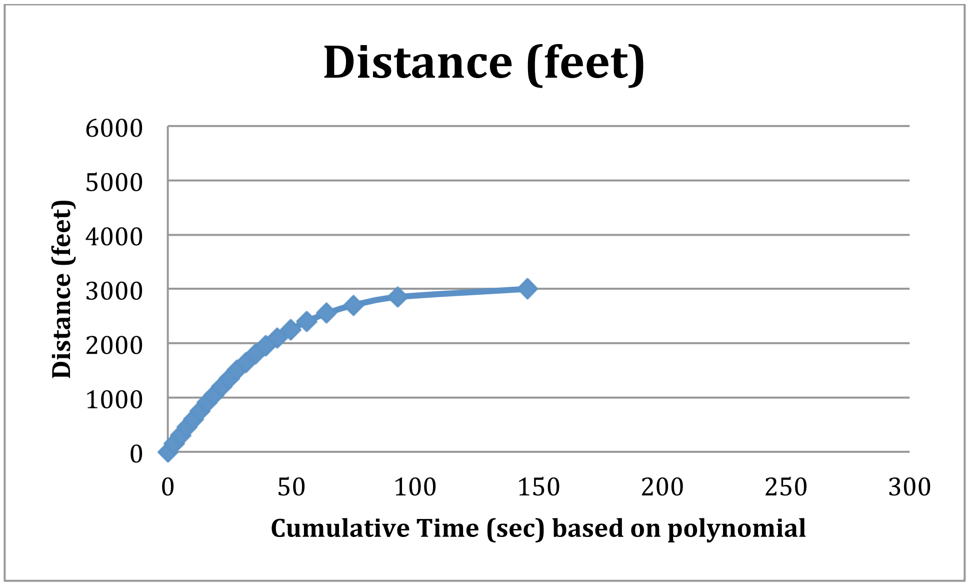

**Figure 4. Train Position as a function of Time After Applying Breaks.**

Finally, we have the train speed as a function of time since brake application.

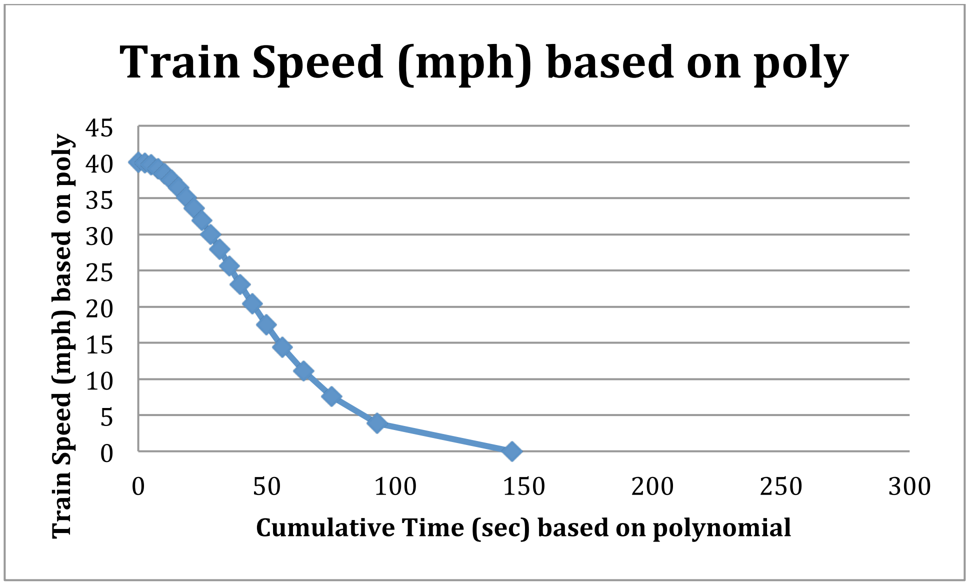

**Figure 5. Train Speed vs. Time Elapsed since First Brake Application.**

In Figure 5, we see that the speed has a point of inflection with respect to time. This is to be expected, because the acceleration is zero by construction before the brakes are applied, zero when the train has come to a stop, and negative in between. Because the acceleration is continuous function of time, it must attain a local minimum between the brake application and stopping instances. The result follows.

The foregoing describes train speed evolution when a service braking curve is in use. Penalty and emergency braking would entail sharper drops in speed upon application of the brakes, affecting performance requirements accordingly.

A comparison of braking curves with initial speeds of 100 mph and 40 mph shows that the predicted braking time is much shorter (~116.3 sec) at the higher speed. This is to be expected, because the train has to brake much more abruptly to stop within the same distance at a higher speed. This implies that a much greater braking force is needed to bring the high speed train to a stop within the same distance if the slow and fast trains have the same mass. Otherwise, the faster train must have much lower mass if the train is to be stopped with the same braking force, because $Fd = mv^2/2$, where $d$ is the stopping distance (fixed) and $F$ is the braking force (also fixed).

It is also worth noting that the average speed with respect to time is not the same as the average speed with respect to distance. The former is the area under the curve in Figure 5 divided by the time for the train to stop. The latter is the area under the curve in Figure 2 divided by the stopping distance, which has a very different shape. Moreover, as the train slows down, it will take larger and larger amounts of time to travel equal distances.

## 4.3. Formulations of QoS and Performance Requirements for Message Delivery

### 4.3.1. Factors Affecting Timely Message Delivery to the OBU from the WIU

Factors affecting timely message delivery from the WIU to the OBU include radio interference, processing delays in which equipment state changes are converted into Wayside status messages within the WIU itself, and propagation delay. Processing delay within the OBU is factor affecting timely action upon Wayside status messages once they have arrived. Our focus here is on QoS requirements for message delivery in a beaconing environment, and on message delivery delays.

### 4.3.2. Requirements for Quality of Message Delivery

#### *4.3.2.1. Beaconing and Forced Train Stops*

A train coming within receiving range of a WIU will either passively receive status messages at regular intervals or receive status messages in response to a *getWIU* request [3]. The status messages contain the status of the signals, switches, grade crossing sensors, track defect sensors, and other wayside devices and the like being monitored by the WIU. The devices notify the WIU directly whenever their status changes.

One rule might be that OBU will treat a wayside status message as stale and invoke the fail-safe stopping procedure if the next wayside status message is more than some time $T_w$ seconds old on arrival. Since clocks at the WIU or BOS and on the OBU are all synchronized using GPS, clock drift should not be severe enough to prevent this stopping rule from working properly We can be reasonably assured of this as long as the requirement that there be no more than $\pm 2000$ msec of drift in an eight hour period is met [S9202]. In addition, one must take into account the possibility that braking on freight trains is binary: it is applied all at once or not at all. Penalty braking is always applied first; emergency braking is only applied if the penalty brake does not bring the train to a speed less than or equal to that given in the braking curve. This in contrast with possibly computer controlled dynamic braking, in which brakes are applied automatically according to the nature of the consist, speed, gradient, and other track conditions, or a more nuanced braking that might be used for passenger trains to ensure platform access from all passenger cars, or at least from a defined subset of them.

Figure 5 shows that the speed vs. time curve corresponding to our synthetic braking curve has negative slope and a point of inflection where the deceleration is maximum. It is desirable for the extension of the signal-based movement authority to be given as early as possible, preferably to the left of the point of inflection, so as to minimize the speed reduction while permission to proceed is en route to the OBU. It is unlikely that permission to proceed would be rescinded for operational reasons, but if it were, a penalty braking curve would have to be invoked as quickly as possible. To ensure timely delivery,

- If permission to proceed is granted after brake application, the message delivery time should be short enough to prevent a slowdown greater than a specified level.
- If permission to proceed is being rescinded or a stop or restrictive signal is otherwise being sent, the corresponding message delivery time should be short enough to prevent the train from traveling a specified distance before it stops.

In either case, this means that the time between beacon transmissions $t_b$ must be less than the maximum chosen value of $t_D - t_0$ . Subject to calculations, the foregoing suggests that the higher the speed of an approaching train, and/or the heavier the train, the shorter the maximum time between beacons must be. If the train is subject to binary braking, we can require that the train receive a status message to stop or proceed at least $T$ seconds before it arrives at the point of automated brake application. If the train is traveling at 40 miles per hour, a delivery time requirement of 10 seconds would result in the train moving (40x10/3600) = 1/9 = 0.1111 miles before commencement of brake application. If the train is subject to a more refined braking policy with a requirement that it lose no more than 5 mph from its initial speed of 40 mph, delivery of a message to proceed should arrive no more than 37 sec after brake application based on the quadratic braking curve shown in Fig 5. Notice that we are specifying a performance requirement in terms of time, not distance. In Fig. 5, we see that the time to reduce the speed from 40 mph to 35 mph is just under 50 seconds. In Fig. 4, we see that the distance traveled in this time is about 2800 feet. Of course, these results would be different if different braking curves were used.

#### *4.3.2.2. QoS Models*

Suppose that a railroad decrees that the maximum permissible time between successful wayside status message arrivals is $T_w$, or that the maximum age of a message is $T_w$, and that the configured time between WIU beacon transmissions is $t_b < T_w$. Then, the maximum number of beacon transmissions that can occur without being received before the brakes are automatically applied is $N_B = \lfloor T_w/t_b \rfloor$. $T_w$ should be chosen with respect to braking curve properties and the stopping requirements discussed in the previous subsection. If there is some probability that beacon messages will be lost, we could assume that the number of beacons before the second one arrives at the train has a geometric distribution with a parameter $p$ equal to the probability of packet loss, regardless of cause. A wayside status message might be lost or corrupted because of radio interference or for some other reason, such as jamming or a message replay attack by a saboteur. All of these factors, and more, could affect the value of $p$. Then,

- The probability that two successive beacon messages are successfully received without interruption is $1 - p$.
- The probability that one beacon message is not received between two successful arrivals is $p(1 - p)$.
- The probability that two successive beacon messages are not received between two successful arrivals is $p^2(1 - p)$.
- The probability that $k < N_B$successive messages are not received between two successful arrivals is $p^k(1 - p)$.
- The probability that the train will be stopped because at least $N_B$ successive messages have failed to arrive between two successful arrivals is one minus the probability that $N_B - 1$ or fewer messages failed to arrive in succession between two successful arrivals. This is equal to

$$p_{STOP} \quad = \quad 1 - (1 - p) \sum_{k=0}^{N_B - 1} p^k$$

$$= \quad 1-(1-p)(1-p^{N_B})/(1-p)$$

$$= \quad p^{N_B} \qquad \textbf{(1)}$$

This expression leads us to a requirement on the probability of a packet being lost. If we require that the probability that the train needlessly come to a stop be less than some small quantity, $\epsilon$, we obtain a constraint on the probability that a packet is lost due to radio interference or other cause:

$$p^{N_B} < \epsilon$$

Or

$$p < \sqrt[N_B]{\epsilon}. \qquad \textbf{(2)}$$

Put another way, if one regards an examination of the stream of wayside status messages as a sequential test that it is unsafe for the train to proceed because the message stream lacks sufficient quality, the probability of getting a false positive should be less than or equal to $\epsilon$. To illustrate, suppose that we require that the probability that a train is stopped unnecessarily because of radio interference be less than $10^{-6}$. Then, if wayside status messages are sent one second apart and the allowed time between successful transmissions is 12 seconds, we have

$$p < \sqrt[12]{10^{-6}}$$

$$\sim \quad 0.3162\ .$$

This shows that, with these configured values, the delivery requirement can be met if just under a third of the messages are lost or damaged in transit. If the maximum permissible age of a wayside status message is 12 seconds and the messages are transmitted every 4 seconds, then $N_B = \frac{12}{4} = 3$, and the corresponding probability of non-delivery of a valid wayside status message must be less than $\sqrt[3]{10^{-6}} = 0.01$. Notice that the required loss probability is highly non-linear with respect to the ratio of the maximum message age to the time between beacon messages. A higher ratio makes the system far more robust, since more .beacon messages can be transmitted during the maximum aging time, provided that the increased number of beacons does not result in message loss because of radio interference or excessive use of bandwidth.

Since the number of consecutive transmissions that are unsuccessful is geometrically distributed with parameter $p$, the mean number of consecutive unsuccessful transmissions is $p/(1-p)$. If $p=0$ we have purely error-free transmission. If $p=1$, every message transmission is unsuccessful, and wayside status messages are hardly getting to their destinations at all, because the expected number of consecutive unsuccessful transmissions is infinite. Interestingly, the expected number of consecutive successful transmissions of beacons is the reciprocal of the expected number of unsuccessful ones, i.e., $(1-p)/p$. Notice that if $p = 1 - p = 1/2$, the expected consecutive number of either successful or unsuccessful beacon transmissions is one, meaning that on average every other message is getting through, as we would expect.

These results show that the I-ETMS beaconing mechanism for implementing PTC is fairly robust in the face of radio interference with the configured transmission interval and timeout value for automatic brake application. Even if the probability of losing any wayside status message is 0.5 (an indicator of severe interference), fewer than three in 10,000 trains will only be erroneously stopped if stopping is triggered by 12 consecutive losses. Normal operations can continue if this level of risk of unnecessary delay is deemed tolerable. Fig. 6 shows the maximum permitted probability of a wayside status message not getting through to the OBU as a function of the probability that a fail/safe stop must be invoked given that the train does not have to stop, for different numbers of successive lost messages between arriving messages.

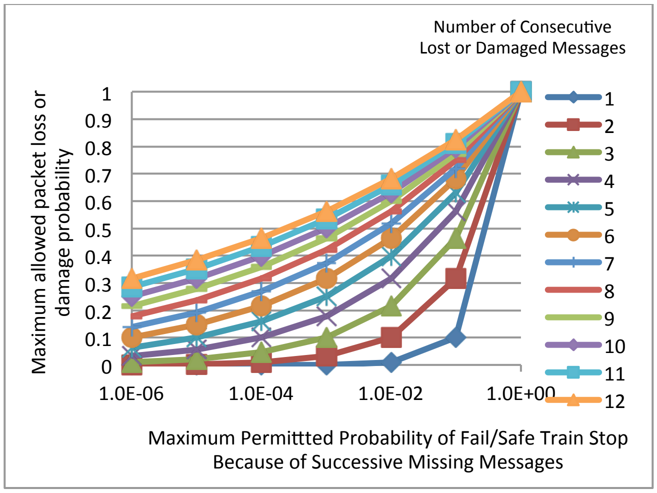

**Figure 6. Maximum permitted probability of lost or damaged wayside status messages for different numbers of permitted successive missing messages between successful message arrivals.**

The plot shows that permitting large numbers of consecutive lost messages enables the maximum allowed probability of losing a message to increase, regardless of the probability that a train is stopped because of consecutive message losses, as might be expected. The degenerate case of allowing erroneous stops without restriction ($p_{STOP} = 1$) means that all status messages can be lost. Restricting the probability of an unnecessary stop to one in a million train movements means that the probability of an undelivered packet should not exceed about 0.32 when stopping occurs after twelve consecutive undelivered beacon messages. When we restrict the number of consecutive undelivered messages to six, the maximum allowed packet loss probability drops to 0.1.

### 4.3.3. Application: Beacon Interval Configuration Guidelines

Equation (12) and Figure 12.6 can be used to guide railroad planners in the configuration of beacon frequency based on known or estimated values for the probability of losing message and the desired maximum probability of a train being stopped on account of wayside status message loss. If the probability of status message loss is high, the risk of erroneously stopping a train because of message loss can be reduced by shortening the interval $t_b$ between beacons. This increases the number of wayside status messages that are sent within time $T_w$. For example, if the probability that a valid wayside status message is not delivered within 12 seconds is approximately 0.3, the odds of an erroneous stop is about one in a million if beacons are sent at intervals of $t_b$=12/12=1 every second, one per thousand if the interval is $t_b$=12/6=2 seconds, and one in ten if the interval is $t_b$=12/2=6 seconds. Notice the choices of $t_b$and $T_w$ also result in a requirement that the OBU be able to process incoming wayside status messages at a rate that is greater than or equal to $1/t_b$ messages per second. If the OBU is built to meet a requirement of processing one wayside status message every $T_C$ seconds, we must either have $T_C < t_b$, or, depending on how the receive buffer is engineered, put a receive buffer in place with a policy of discarding all but the most recent WIU message.

### 4.3.4. Delay Requirements for Message Delivery

We can use the numerical method suggested by equation (6) to compute message delivery and other delay requirements for a train approaching a WIU that guards an entrance to PTC territory. Since the train is PTC-enabled, it will know of the existence and position of the WIU and either start listening for its status beacon or issue a *getWIUstatus* request once it comes within range of the WIU's radio transmitter. Once the locomotive is in range, the following events occur:

1. The locomotive authenticates the identity of the WIU, taking time $t_a$ to do so.
    a. Note: if HMAC is the sole means of authentication, as will be the case in I-ETMS, this is done for every message. The cost is part of the message processing time. In that case, $t_a = 0$.
2. Either the WIU sends a beacon at regular intervals continuously, or the locomotive must issue a *getWIUstatus* request to start the beacon. The time taken to receive a beacon is the expected time between successful beacon transmissions, $t_S$, together with the time for the WIU to respond to a *getWIUstatus* request, $t_g$. If the WIU is beaconing steadily without awaiting a *BeaconRequest* message, we set $t_g = 0$ without loss of generality.
3. Either the engine engineer takes time $t_R$ to react to the display on the OBU HMI or wayside display of the signal, or, if the signal is at danger, the OBU starts activating the brakes according to the braking curve if the engineer has ignored it, taking time $t_{OBU}$ to do so. U.S. government regulations stated in 49 CFR 236.563 and 49 CFR 236.831 state that if the train is required to stop, the time elapsed from when the onboard apparatus detects a more restrictive indication until brake application commences shall not exceed 8 seconds. Hence, we have $t_{OBU} < 8$ seconds [13].

For a given probability of failed beacon transmission $p$, the expected time to receive a successful beacon is approximately the expected time between successful transmissions, or, equivalently, the expected number of failed transmissions multiplied by the beacon interval time, plus the expected time to the next beacon emission, which is $t_b/2$. Thus, we have

$$t_S = \frac{t_b p}{1-p} + \frac{t_b}{2}$$

$$= \frac{(1+p)t_b}{2(1-p)} \qquad \textbf{(3)}$$

Notice that with $p = 0$, the expected time to successfully receive a wayside status message is $t_b/2$, and that with $p = 1$, the expected time is infinite, because all WIU transmissions are failing, as one would intuitively expect.

Let $t_A = t_{OBU} + \mathrm{t_S} + \mathrm{t_g}$ for notational convenience. Suppose that the track distance from the train's initial point of PTC radio reception to the point at which brakes are applied is based on calculations from the braking curve, and that the train is traveling with constant speed *v*. If the braking curve is followed and the train is to be stopped before the signal of interest, we must have:

$$v(t_a + t_A + t_R) < B \qquad \textbf{(4)}$$

If $x_R = vt_R$ is the distance travelled during the engineer's reaction time, we then obtain the following upper bound on the time for a locomotive to be authenticated by the back office server:

$$\boxed{t_a + t_A < \frac{(B - x_R)}{v}, x_R < B} \qquad \textbf{(5)}$$

This equation shows that a long reaction time shortens the maximum permissible time to activate the brakes according to the braking curve, and that the maximum allowable activation time is inversely proportional to the speed of the train. Moreover, the reaction distance must be less than the distance to the point at which the braking curve is activated. Interpreted another way, if the brake activation time cannot be reduced, then either the speed *v* in the approach zone to the PTC-controlled track must be decreased accordingly, or the braking distance *B* must be increased accordingly. Combining equations (12), (13), and (15) gives us a relationship between successful transmission probability, the desired probability of the train wrongly being brought to a stop, speed, and braking distance, because the definition of $t_A$ implies that

$$t_a + t_{OBU} + \mathrm{t_S} + \mathrm{t_g} < \frac{(B - x_R)}{v}, x_R < B$$

whence

$$t_a + t_{OBU} + \frac{\left(1 + \sqrt[N_B]{\epsilon}\right) t_b}{2(1 - \sqrt[N_B]{\epsilon})} + \mathrm{t_g} < \frac{(B - x_R)}{v}, x_R < B \qquad \textbf{(6)}$$

If this inequality cannot be satisfied by a reasonable set of parameters, the train cannot be brought to a stop within the requisite distance while keeping the probability of an erroneous stop below the desired level.

If the train is eligible to proceed in more permissive manner than is being enforced by the PTC mechanism, it is desirable that the OBU be notified of this as early as possible so that enforcement can be rescinded. This minimizes the adverse effect of unnecessary braking. Referring to Figure 3 , we see that it is desirable to rescind enforcement as soon after application as possible, preferably at the earliest point at which the absolute value of the intended acceleration (the slope of the speed with respect to time) is as small as possible. Put another way, we wish to minimize the loss of kinetic energy due to unnecessary braking by setting the maximum desired message delivery time $\mathrm{t_D} - \mathrm{t_0}$to ensure that the change in kinetic energy, $\frac{1}{2}\mathrm{m}(\mathrm{v_0^2} - \mathrm{v_D^2})$, is as small as possible, where $\mathrm{t_0}$ is the time at which monitoring for brake enforcement starts, $\mathrm{t_D}$ is the time at which notification occurs that no restriction is in place, and the $\mathrm{v_i}$s are the corresponding train speeds, for $\mathrm{i} = 0, \mathrm{D}$ and $\mathrm{v_0} > \mathrm{v_D}$. For the requirement to be unambiguous, we must specify how large we would allow $\mathrm{v_0} - \mathrm{v_D}$ to be, and then derive $\mathrm{t_D} - \mathrm{t_0}$ numerically using equation (6).

## 4.4. Application

One can apply the methods described here to any form of braking curve that is continuous. Some modifications to the numerical method for computing braking times would be needed if the braking curve is only piecewise differentiable (i.e., it has corners in it, but no jumps). These conditions appear to hold for the braking curves we have seen in the published literature. We have also shown how these performance requirements are related to message transmission quality when using a beaconing mechanism. A requirement for a low probability of stopping the train solely because of poor beacon transmission quality and knowledge of the probability of beacon message being delivered (or not) can be used to determine the necessary time interval between beacon transmissions once the stopping distance and speed have been specified. Because stopping a train is costly, reducing the risk of unnecessarily doing so will have a positive economic impact on railroad operations while maintaining a proper standard of safety.

We can use the numerical method suggested by equation (0.6) to compute delay requirements for a train approaching a WIU that guards an entrance to PTC territory. Since the train is PTC-enabled, it will know of the existence and position of the WIU and either start listening for its status beacon or issue a *getWIUstatus* request once it comes within range of the WIU's radio transmitter. Once the locomotive is in range, the following events occur:

1. The locomotive authenticates the identity of the WIU, taking time $t_a$ to do so.
2. The locomotive authenticates itself with the BOS if it is within range of it, taking $t_{BOS}$ to do so.

3. Either the WIU sends a beacon at regular intervals, or the locomotive issues a *getWIUstatus* request. The time taken to do so is either the maximum time between beacon emissions,$t_b$, or the time for the WIU to respond to a *getWIUstatus* request, $t_g$.
4. Either the engine driver takes time $t_R$ to react to the in-cab or wayside display of the signal, or, if the signal is at danger, the OBU starts activating the brakes according to the braking curve if the driver had ignored it.

Let $t_A = \max(t_b, t_g)$ for notational convenience. Suppose that the track distance from the train's initial point of PTC radio reception to the point at which a braking curve must be applied is $B$, and that the train is traveling with constant speed $v$. If the braking curve is activated and the train is to be stopped before the signal of interest, we must have:

$$v(t_a + t_A + t_R) < B$$

**(0.12)**

If $x_R = vt_R$ is the distance travelled during the driver's reaction time, during which the speed of the train is assumed constant, we then obtain the following upper bound on the combined time for a locomotive to be authenticated by the back office server and the time to be authenticated by the WIU:

$$t_a + t_A < \frac{(B - x_R)}{v}, x_R < B$$

**(0.13)**

This equation shows that a long reaction time shortens the maximum permissible time until the train's speed must follow the braking curve, and that the maximum allowable activation time is inversely proportional to the speed of the train. Moreover, the reaction distance must be less than the distance to the point at which the braking curve is followed. Interpreted another way, if the brake activation time cannot be reduced, then either the speed $v$ in the approach zone to the PTC-controlled track must be decreased accordingly, or the braking distance $B$ must be increased accordingly.

## 4.5. Potential Impact on Engineering and Operations

The foregoing results show that braking properties, requirements for the quality of beacon transmission, and the desired probability of erroneously stopping a train are closely related. The impact of this relationship is that transmission quality cannot be engineered independently of service and braking properties, or vice versa. Competing stakeholders and those who coordinate their activities will need to weigh tradeoffs and the cost of engineering braking capability to accommodate radio interference. The regulatory and cost factors associated with these tradeoffs are outside the scope of this paper. Here, we devised a means of analyzing the technical aspects of the relationship that will help inform choices of tradeoffs.

## 4.6. Factors Affecting the WIU Load

The load on the wayside interface unit (WIU) is determined by the frequency with which trains interact with it and the number of wayside devices (i.e., hazard warning devices, signals, and switches) with which the WIU must interact for every train movement. The number of wayside devices connected to a WIU could be quite large. Alstom has recently announced that Amtrak will acquire 250 WIUs for its Northeast Corridor (NEC) line [Alstom2011]. Since this line has multiple parallel tracks in several places between Boston MA and Washington DC, it is clear that there is more than one device per WIU.

### 4.6.1. The Need for Guidelines to Determine the Number of Wayside Devices for a WIU

Guidelines are needed to ensure that no WIU will be saturated because it interfaces to a large number of devices and often receives *getWIUstatus* messages from trains. In practice, the WIU will broadcast WIU status message continuously in heavily traveled areas. There, the limiting factor of WIU performance is the number of updates it receives directly from the wayside devices that are attached to it. . As we shall see below, the communications protocol between the Wayside, Locomotive, and Office Segments only loosely restricts the number of devices that interface to a WIU.

### 4.6.2. Activities Contributing to WIU Load

Every train movement through a block involves the following activities:

1. Authentication of the train by the WIU if the train sends a *getWIUstatusmessage.* In I-ETMS, this is done solely with HMAC.
2. Authentication of every WIU status message by the train. IN I-ETMS, this is done solely with the HMAC field.
3. Setting the field in the WIU status message corresponding a signal that is set at danger once the train passes it once the train has entered it,
4. Setting the field in the WIU status message corresponding to the signal that becomes more permissive once the train leaves the block the signal controls.

Without loss of generality, suppose that there are $k_i$ objects in the $i$th of $L$ interlockings, for $1 \leq i \leq L$. For modeling purposes, we treat a block of track without switches guarded by a signal as an interlocking with only one object. Suppose further that interlocking $i$ has $\lambda_i$ trains coming through it in unit time during the peak hour, or the designated peak period of shorter duration (e.g., 5 minutes).[1] Then, the WIU should be engineered to support $\lambda_i$ authentications and $2\,\lambda_i k_i$ interlocking changes in unit time, where $k_i$ is the number of devices that must be changed in the $i$th interlocking. By "support," we mean

(a) that no resource within the WIU will have more than 50% utilization, including bandwidth in and out,
(b) that authentication and interlocking communication delays with signals and the BOS will have a designated average value or less, and that 95% individual delays will be below some chosen upper bound.

[1] Perhaps engineering for the peak 5 minutes of the day is appropriate, because some central stations may have several movements close to the hour.

### 4.6.3. WIU Status Bits

At the time of writing, we are not aware of any guidelines for capping the number of devices connected to a single WIU. The maximum number that could be connected to a single WIU is not severely restricted by the number of status bits available in packets to be sent between the OBU and the WIU.

In Section 2.1 of [S9352B], it is stated that there are 1,944 bits assigned for the device status array. The number of bits for each type of device is shown in table 12.1. Table 1shows the device code conventions. We have copied it from [S9352B] directly.

**Table 1: Device Status Code Conventions**

| Device Type | Status Code Length (bits) | Value (MSB … LSB) | Indication |
|---|---|---|---|
| Switch | 2 | 0b00 | Switch Position INDETERMINATE |
| | | 0b01 | Switch Position REVERSE |
| | | 0b10 | Switch Position NORMAL |
| | | 0b11 | Switch Position Indication ERROR |
| Signal | 5 | * | Implementation specific. |
| Hazard Detector | 1 | 0b1 | No Condition detected |
| | | 0b0 | Condition detected |

Let P denote the number of switches (called Points in the UK), S denote the number of signals (presumably home signals, because distant signals show the same status), and H denote the number of attached hazard detectors. P, S, and H are all non-negative integers. Because we have 2 bits for a switch, 5 for a signal, and 1 for a hazard detector, the combination of the numbers of devices of each type must satisfy the following inequality:

$$2\mathrm{P} + 5\mathrm{S} + \mathrm{H} \leq 1944 \tag{0.14}$$

Geometrically, the space of possible combinations is a set of points bounded by a plane with normal (2,5,1) having a distance of 1,944 from the origin, as well as by three mutually perpendicular planes, each passing through two of the P, S, and H axes.

From this, we see that the layout of the device status array in WIU status messages is not a tight constraint on the number of devices attached to each WIU. Rather, it appears that the dominant constraint will be the processing capacity of the WIU and possibly the bandwidth capacity. We need a better understanding of the demands that will be made on the WIU to formulate a performance requirement.

### 4.7. Restriction on the Number of Signal Aspects and Status Codes

The signal status code is 5 bits wide. If there are $a$ aspects and $c$ error conditions, we have the constraints

$$\begin{gathered} a + c \leq 32 \\ a > 0 \\ c > 0 \end{gathered} \qquad (0.15)$$

This does not appear to place a severe limit on the number of signal aspects that can be used.

### 4.8. Track Layouts, Timetables, and Throughput

Clearly, number of train movements through, into, or out of a sector controlled by a BOS affects the demand placed on the resources of the PTC system through direct interactions between trains and WIUs. Similarly, the load on the WIUs is affected by train movements and the number of devices that interface with each WIU. Additional load on the system, including the associated network bandwidth, occurs because WIUs, OBUs, and BOSs are synchronized via messages transmitted at regular intervals known as heartbeats. Failure to deliver heartbeats in a timely manner could result in service disruptions, because the heartbeats are used to verify that corresponding segments are up and running and communicating with one another. The frequency of interactions between trains and WIUs depends not only on the frequency of train movements, but also on how often trains pass signals and switches attached to a given WIU. In particular, delay requirements depend on the lengths of blocks and how fast trains travel through them, as well as on the distances between WIUs and the frequency with which trains approach them.

The timetable describes the frequency of train movements from particular directions, while the track maps we have been given for Trenton and its southern approaches show how many wayside devices are traversed each time a train goes through an interlocking. Because the maps are not recent, they do not show how WIUs will be deployed and which devices will be controlled by them. This would be useful information to have, so that we can devise guidelines for the loads the WIUs are meant to sustain.

## Section 5. Bandwidth Usage Between OBU and BOS

### 5.1. Overview

In the I-ETMS, the set of communication messages between the railroads' Back Office Servers (BOSs) and the trains' On-Board Units (OBUs) is given in draft Interface Control Document (ICD) S-9352A [S9352A]. ICD S-9352A does not mention any challenge/response or password-based authentication messages for setting up a connection between the train and the back office server. The word “authentication” appears just once in [S9352A], with reference to HMAC keys. This suggests that I-ETMS relies entirely on confidence in the origin of the data for authentication, using HMAC as described elsewhere in this report. In this section, we model bandwidth usage based on our understanding of the message traffic that might ensue between the OBU and BOS given the current status of [S9352A]. We also examine the delays that might be experienced by hypothetical explicit authentication messages were they to become part of the S9352A message set.

Part of our original brief was to devise a framework for predicting authentication delays when trains cross railroad boundaries. The back office servers of each railroad are concerned with the ongoing monitoring of train status as well as with controlling handovers across network boundaries. An architecture document [S9001] indicates that the handling of both will be done on the same hosts with messages using the same bandwidth for communication between trains and back office servers.

We do not know if the use cases, operational conditions for the dispatch of messages between the train and back office, and how often they are sent have been defined by the railroads. If they have, the definitions are proprietary to the railroads and have not been made public as of the time of writing

The absence of a publicly available description of the conditions for sending messages between the train and the back office server of any type necessitates the use of assumptions about them to build models to estimate (a) the message handling delay and (b) the bandwidth that will be consumed by communications between a train’s On-Board Unit (OBU) and the railroads’ back office servers (BOSs). The basis for these assumptions is laid out in Section 0.

Since explicit authentication messages are not part of [S9352A], we model authentication delays on the assumption that the message set would be augmented to include them. The goal of doing so is to provide insight into the factors affecting authentication message delays per our original brief, should explicit authentication eventually be included. Performance models are needed that predict bandwidth usage and the response times associated with handling messages, including those associated with an authentication scheme based on message exchange. Such messages would augment the present scheme, which relies solely on the mutual recognition of the contents of HMAC fields for authentication. Because HMAC fields could be forged in a replay attack as discussed elsewhere in this report, it may not be sufficient to rely on them for authentication in their present form.

Our modeling approach is as follows:

1. We use our understanding of operations to choose the number of times each train will send or be sent a particular type of message per hour.
2. We estimate the number of trains that are present in proximity to a radio tower and the number of trains that are in contact with a particular BOS at any time,

and use that information to estimate the respective message volumes and bandwidth utilizations.

3. We use coarse models of each BOS and guesses about traffic patterns between trains and BOSs to make rough predictions about message volumes and the resulting processing delays in handling authentication messages or any kind of message in the [S9352A] message set.
4. We use a coarse model to estimate the channel utilization between each OBU and the radio tower nearest to it. We do not consider bandwidth usage between the radio towers and the BOSs in detail because this is not seen to be a constraining factor unless there is heavy traffic denoted to position notification.
5. We vary the parameters of these coarse models to assess the impact of different operational procedures on traffic patterns and on resource utilizations. This will help to identify potential limiting factors on resource usage.

## 5.2. Messages from BOS to OBU and Vice Versa

The Interface Control Document (ICD) [S9352A] describing the set of communications between the OBU and the BOS defines 47 types of messages that may be sent from the BOS to the OBU and 47 other types of messages that may be sent from the OBU to the BOS. The numbers of message types in each direction were not equal in earlier versions of this ICD. Hence, it should not be assumed that there is a one-to-one correspondence between the messages sent in one direction and the messages sent in the other.

While many of the types in one direction have a corresponding type in the other direction, there is not a one-to-one correspondence between messages sent one way and messages sent the other. For example, a BOS's request for position notification may contain a specification of the criteria an OBU shall use to send repeated notifications, based on distance traveled or the time since the last message, both of which may be configured. In addition, position notifications are sent by a train when it begins to move from rest, when it stops, when it crosses a switch, when it passes a signal, and when it traverses a grade crossing. Some types of messages will usually only be sent to the train when the train is stopped in a station or in the yard, e.g., the messages containing the on-board track data base and those announcing the results of pre-departure tests. Movement authorities are sent as needed or as requested, as are fault messages and notifications that the OBU has engaged penalty or emergency braking because the driver passed a stop signal or exceeded the speed limit. How often these types of messages are sent depends on driver behaviour, how fast the train is moving, how many switches it traverses per hour, and other factors that may be particular to the train's current location. In the absence of operational guidelines and historical data, we must make assumptions about the message rates in each direction under various conditions.

Our analysis suggests that Position Notification Messages from the OBU to the BOS could be the ones most frequently sent if the BOS asked for them. The ICD and other draft standards documents are silent on the reasons for choosing that messages be sent based on distance traveled or based on fixed time intervals, or on whether position notification is required at all. Each railroad may decide this on its own. How often and when messages should be sent depends on how the data will be used. In the European Train Management

System, position notification messages are only sent every 15 seconds [ERTMSCITATION], because track circuits are extensively used to provide location information. At the AMTRAK control center in Wilmington, Delaware, train positions are only indicated to the nearest block [Bondi2012]. This is in keeping with traditional railway practice [Pachl2002]. At least one U.S. railroad would like to be able to identify the position of all of its trains with greater precision if complete radio coverage were available [Murphy2012].

We note the following considerations and concerns regarding train position notification:

1. Trains have neither constant length nor constant speed. Trains do not have constant length because they are longer when a train is going uphill, shorter when going downhill, longer when accelerating if being pulled, and become shorter when pushed uphill from rest if the locomotive is at the rear end of the train instead of at the head end, because coaches must be pushed closer together for the train to start moving. Therefore, one can only be certain that a train has cleared a switch if proper notification is sent when both the head end and the rear end of the train have cleared the switch.
2. Sufficiently accurate position notification might be used to manage train dispatching and the granting of authorities if there are no track circuits. This is especially likely in dark territory, where there are no signals. Dark territory is lightly traveled, so the greater concern is radio signal strength rather than bandwidth congestion. In a report to Congress, the Federal Railroad Administration mentions the possibility of a required precision for position location of one ten thousandth of a mile, which works out to 5,280/10,000 = 0.5280 feet, or 6.336 inches (about 16.1 cm) [FRA2012]. The reason for such precision is not stated. We speculate that it could be due to the need to align side doors at the end of the train with platforms.
3. Position notification could use a large fraction of the bandwidth if it is configured to occur frequently. In congested areas, track circuits are more likely to be used for position notification than in uncongested areas, so position notifications would be needed there less frequently than in uncongested areas, such as in dark territory that might not be equipped with track circuits.
4. The ICD [S9352A] lists a message to be sent by the BOS that requests the configuration of the OBU to send position notification messages either at constant time intervals or at constant travel distances, in addition to when the train passes signals or switches. If the train is moving slowly, messages sent at constant distances or when the train crosses switches or goes past signals will occur further apart in time. If the train is moving quickly, it will move further during constant time intervals than if it were moving slowly, and messages sent when traversing switches will be sent more frequently. If fine time granularity for position notifications is required, e.g., because of braking curve considerations, the position messages should be sent at short time intervals. As a train moves from a congested region to an uncongested one, the BOS can send a message to the OBU requesting position notifications at shorter time intervals or at shorter distances.
5. Since the position indicated by GPS is only updated once per second [S9352C], there is no point in sending position notifications at time intervals of less than one second. This limits the bandwidth utilization due to position notification messages.

Enforcement notifications are sent automatically from the OBU to the BOS whenever the OBU intervenes to apply the brakes on a locomotive that is exceeding the speed limit or whenever the OBU applies the brakes after a stop signal has been passed. A message is also sent whenever there is a train handling exception. These messages are sent over the 220 MHz spectrum when the train is en route, as opposed to being within range of a WiFi network in a station or yard.

## 5.3.Bandwidth Usage Calculation

Two components of bandwidth usage need to be considered: the division of the bandwidth into channels, and the utilization of each channel on its own. The number of channels available depends on how the frequency range has been broken up. The utilization of the channel pool depends on the duration of a connection between each train's OBU and the BOS and the rate at which channel connection requests occur. The utilization of each channel on its own is the ratio of the rate at which bits are offered for transmission to the bit rate that can be offered by the channel.

Under FCC rules, a 25 kHz channel must have a minimum available bit rate of 19.2 kbps when transmitting in the range 412 MHz-512 MHz [FCCapplicationHint2012] CAN WE OBTAIN A BETTER CITATION THAN THIS?. For modeling purposes, we assume that this is the minimum available bit rate for a channel in the 220 MHz spectrum available for PTC. We assume that a 5 kHz channel would have one fifth of this bandwidth at least, i.e. 3.84 kbps = 19.2/5 kbps. The available bit rate is obtained using modulation and encoding schemes that are beyond the scope of this paper. We compute the bandwidth utilization per train from BOS to OBU and vice versa by adding up the offered rates at which different messages occur, multiplying by the average packet size, and then dividing by the available bandwidth per channel. Multiplying this quantity by the average number of trains in range of a radio mast yields the total normalized demand for bandwidth near that mast. Multiplying this quantity by the average number of trains corresponding with a BOS yields the total demand for bandwidth by all trains summed over all radio zones with which the BOS communicates. We consider bandwidth usage en route and in the yard separately

Let us assume that communication medium $j$ has effective bit rate per channel $\mu_j$, and that the size of message type $m$, including the EMP header, is $\sigma_m$ bits. We let $j = W$ for WiFi and $j = Q$ for 220 MHz radio. Denote the rate at which a particular train generates messages of type $m$ when communicating with BOS $b$ by $\lambda_{b,m}$. We assume for the convenience that a particular message type is transmitted only in the yard or only en route, but not both, and that it is always transmitted in one direction (BOS to train or train to BOS), but not the other. We denote the set of messages sent in direction $d$ at over medium $j$ by $M_{d,j}$where $d \in \{TTB, BTT\}$ and $j \in \{W, Q\}$. The channel bandwidth utilization involving BOS $b$ in direction $d$ is given by

$$\alpha_{bdj} = \sum_{m \in M_{d,j}} \lambda_{bm}\sigma_m/\mu_j \tag{0.1}$$

For minimum transmission delay, we desire that $\alpha_{bdj}$be as small as possible. Since this refers to utilization per channel, we require that

$$\alpha_{bdj} < 1 \tag{0.2}$$

The arrival rates used in equation (0.1) must be derived from our understanding of operational conditions. The packet sizes are derived from draft standards documents. A worst case analysis may be performed by setting all of the values of $\sigma_m$ to their largest possible value, and the effective bit rate $\mu_j$to its smallest possible value. This gives us an upper bound on the bandwidth utilization for a given combination of message arrival rates. The values of the arrival rates depend on operating conditions. For traffic from train to BOS while the train is en route, the message types with the potentially largest rates are position notification messages and train status messages. For messages from BOS to train while the train is in the yard or station, the largest possible messages contain mandatory directives, the train consist, and the contents of the on-board track database. These are all iniatialised before every journey. Mandatory directives may also be sent to the train while it is en route if conditions change.

## 5.4. Crossing a Network Boundary

### 5.4.1. Taffic Possibilities and Demand

We consider the following possibilities:

1. A train in its home network authenticates itself to its home network. A train belonging to Carrier A, traveling in network A, authenticates with BOS_A. If the train is traveling to network B, BOS_A queries BOS_B to obtain permission for the train to enter network B.
2. A train belonging to Carrier B, whose crew is employed by Carrier B, is traveling in network A and authenticates with BOS_A. BOS_A then queries BOS_B about the train's identity. If the train is traveling to network B, BOS_A queries BOS_B to obtain permission for the train to enter network B.

We model these possibilities by treating each BOS as a black box server with FCFS queueing. Each of the sequences above corresponds to a chain of visits or interactions with one of the servers. The authentication delay is the sum of the delays visiting each of the servers, together with any network delays over the air and on terrestrial links. The flows for cases 1 and 2 are identical.

### 5.4.2. Queueing Models of the Back Office Servers

The queueing models show how the paths of the high level flows cause demands to be made of the individual back office servers. We regard each BOS as a single synthetic server that is subject to three different types of demands for processing:

- Demand related to intranetwork activity, related to trains that are not crossing a network boundary. We call this intranetwork demand.
- Demand related to a train leaving the BOS's network for a neighbouring network. We call this the outbound demand.
- Demand related to a train entering the BOS's network from a neighbouring network. We call this the inbound demand.

Notice that the outbound demand related to a train moving to a neighbouring network corresponds to the inbound demand at the receiving network's BOS.

The processing effort and delay for authentication under the sole control of either the host network's BOS or the destination network's BOS is about the same. This is illustrated in Figure 7 and Figure 8. Where both a BOS and a train trigger queries to a corresponding network, the processing burden at each BOS will be somewhat larger.

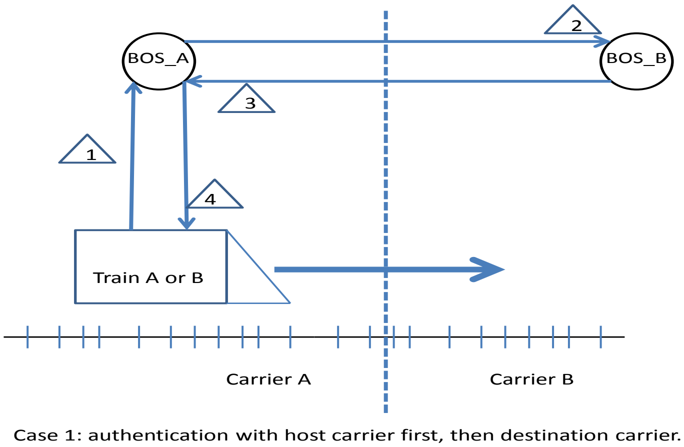
2
BOS_A
BOS_B
3
1
4
Train A or B
Carrier A
Carrier B
Case 1: authentication with host carrier first, then destination carrier.

**Figure 7. High level flow of authentication begun at the host carrier's BOS.**

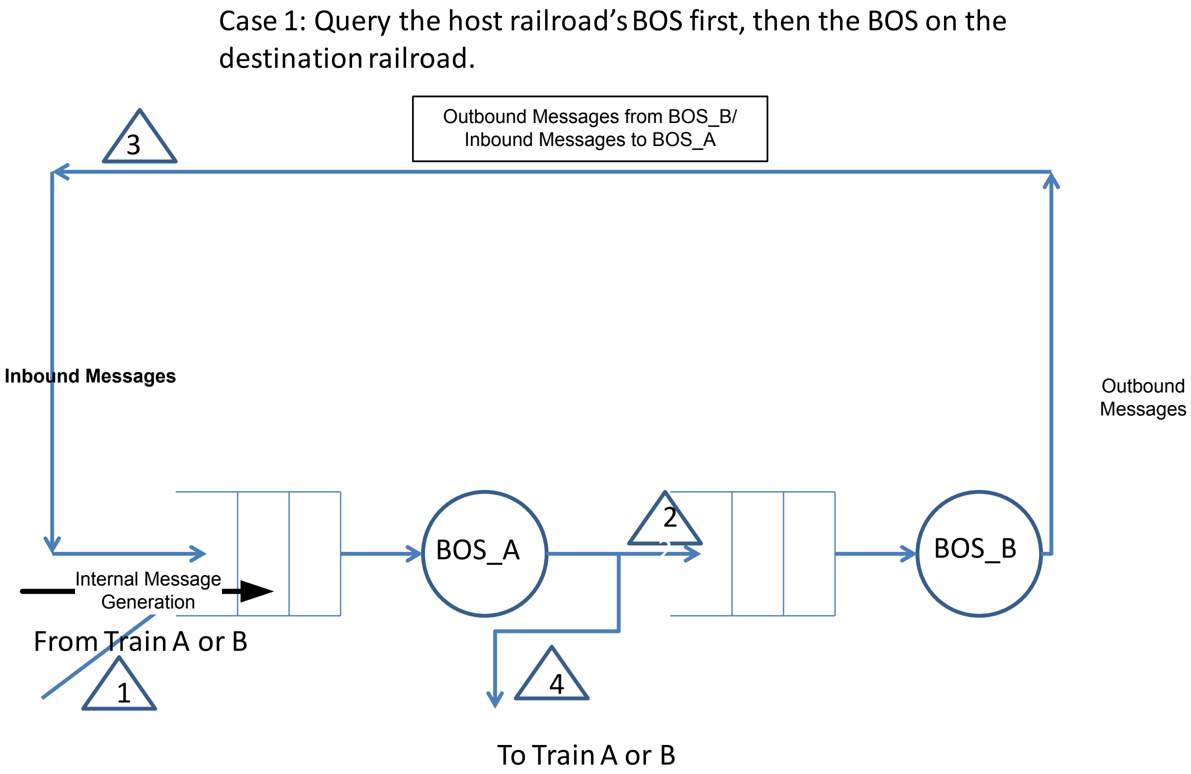
Case 1: Query the host railroad's BOS first, then the BOS on the destination railroad.
Outbound Messages from BOS_B/
Inbound Messages to BOS_A
3
Inbound Messages
Outbound Messages
BOS_A
2
BOS_B
Internal Message Generation
From Train A or B
1
4
To Train A or B

**Figure 8. High level queueing model of authentication begun at the host network's BOS, showing cross traffic.**

Let us consider a set of model parameters for **Error! Reference source not found.**. For simplicity, we model each BOS as a single server fed by messages from trains within each BOS's network, and by messages coming from BOSs in other networks. In this case, we assume that there are only two rail networks, and that each has trains moving from one to the other. Each single server represents the entire set of devices within each BOS. We could model each BOS as a network of queues, but we choose not to do so to simplify our analysis and because we do not have any information about the various devices or the demands that might be made upon them. For trains traveling from B to A, the pattern is the same with _A exchanged with _B. There are trains moving on the A and B railroads at the same, and from A to B and B to A at the same time. Only the A->B message traffic is depicted here Multiple messages may be processed concurrently within each BOS. Messages arrive from the trains asynchronously. Messages may be generated from within the BOS without regard to train movements as well as in response to requests from trains. The ability to process multiple messages concurrently might arise from the presence of multiple threads, multiple I/O devices, and/or multiple processors. Absent detailed information, we treat the server representing each BOS as having the processor sharing discipline (PS), which is essentially equivalent to time slicing in the limit as the duration of the slice tends to zero [BCMP1975]. By the BCMP Theorem, the processor sharing assumption also allows us to avoid making any assumptions about processing time variability. The equations for the mean queue length and the mean response time with Poisson arrivals are the same as those for an M/M/1 queue. The traffic intensity at each queue is the sum of the products of the arrival rates and their corresponding service demands. The service demands are each described by the visit ratio of the message handling type (inbound, outbound, and intranet work) and the mean processing time. Here, the mean processing time is a contrived quantity designed to illustrate the effects of having different fractions of the message activity corresponding to intranet work train movements (staying within the network) and cross border train handling (outbound). If messages are not lost between the BOSs, the outbound traffic sent to one BOS by another contributes a corresponding amount of inbound demand at the receiving BOS. Referring to **Error! Reference source not found.**, we can specify a visit ratio to BOS_A for each of the outbound, intranet work, and inbound components of the work that BOS_A does. Since there is one unit of BOS_A's inbound work to be done for every outbound unit of BOS_B's work, B's outbound visit ratio is A's inbound visit ratio, and vice versa. The visit ratio is the number of times an activity occurs for each message coming from a train or being sent to a train. We lump the total rate of message activities per train corresponding directly with a BOS into one arrival rate. We have chosen arbitrary service times for intranetwork, inbound, and outbound activities for the sake of illustration only. In addition to the service demands, we specify the probability that any message activity is related to traffic that is somehow connected to the other BOS. We vary this probability to reflect the possibility that traffic from network A to network B may be heavier at different times of the day. For instance, trains headed to New York via Amtrak from NJ Transit's North Jersey Coast Line (NJC) run three or four times an hour on weekday mornings, but only once an hour after 09:15. Similarly, trains from New York branch to the North Jersey Coast Line three or four times an hour from 17:00 to about 19:00, but only once an hour the rest of the day. This affects the fraction of traffic that is not crossing a railroad boundary. We expect that $\beta_{NJC,AMTRAK}$ will be larger in the morning rush hour than at other times. Similarly, we expect that $\beta_{AMTRAK,NJC}$ will be larger in the evening rush hour than at other times. We would never expect that $\beta_{NJC,AMTRAK} = \beta_{AMTRAK,NJC}$ at any time of the day, because AMTRAK and NJ Transit combined have far more trains between Newark Penn and Trenton than run between Newark Penn and the NJC.

Parameters for symmetric traffic are shown in Table 1. Asymmetry would be reflected in changes to the numbers of trains present, the numbers of messages of each type sent per train, and the fraction of messages related to border crossings, $\beta_{a,b}$.

**Table 1. Example model parameters for Case 1.**

| BOS | Cross border Outbound | Internal Traffic | Cross border Inbound | | | |
|---|---|---|---|---|---|---|

|  | *Visit Ratio* | *Service Time (sec)* | *Demand (sec)* | *Visit Ratio* | *Service Time (sec)* | *Demand (sec)* | *Visit Ratio* | *Service Time (sec)* | *Demand (sec)* | *Fraction of local message traffic that is bound for the other BOS* $\beta_{a,b}$ | *Messages per hour per train* $\lambda_{b,t}$ | *Average number of trains present* $\bar{N}_b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BOS_A | 1 | 0.01 | 0.01 | 1 | 0.02 | 0.02 | 1 | 0.015 | 0.015 | 0.1 | 689.5 | 125 |
| BOS_B | 1 | 0.01 | 0.01 | 1 | 0.02 | 0.02 | 1 | 0.015 | 0.015 | 0.1 | 689.5 | 125 |

To model the demands on the Back Office Servers, we need estimates of the rates at which the messages of different types occur for each train, and a way to compute the number of trains sending message to a BOS. The total work per train due to messages at the BOS is the sum of the rate at which messages are generated by the BOS itself for each train, and the rate at which messages are sent by each train to the BOS. Using the subscripts TTB and BTT to denote traffic from Train to BOS and BOS to Train respectively, the total message traffic at the BOS per train is given by

$$\lambda_{b,t} = \lambda_{b,TTB} + \lambda_{b,BTT}$$

We have included the index $b$ in this expression because the rates on the right hand side depend on the operating procedures of each railroad, and so may differ from one railroad to another.

The average number of trains communicating with BOS_b, $\bar{N}_b$ , is the average duration of a train journey in the network covered by that BOS, $T_{b,J}$ , multiplied by the number of times that such a journey begins per unit of time, $\gamma_b$. Thus, we have

$$\bar{N}_b = \gamma_b T_{b,J}$$

Hence, the total message rate at BOS_$b$ is given by

$$\Lambda_b = \lambda_{b,t}\bar{N}_b$$

We now show how to combine the arrival rates with the quantities in Table 1. We first introduce some notation.

- The visit ratio for traffic of type $k$ at BOS_$b$ is denoted by $V_{b,k}$, for $b = A, B$ and $k = CBO, CBI, INT$ where $CBO$ denotes cross border outbound traffic, $CBI$ denotes cross border outbound traffic, and $INT$ denotes intranetwork traffic due to trains moving within the network only.
- The corresponding service time per visit is denoted by $S_{b,k}$ and the demand by $D_{b,k} = V_{b,k}S_{b,k}$.
- The probability that a message at BOS_$a$ triggers work at BOS_$b$ is denoted by $\beta_{a,b}$ for $a, b \in \{A, B\}, a \neq b$.

For simplicity, absent further information, we assume that all message processing traffic falls into one of the categories *CBO, CBI,* or *INT.* Our modeling approach reduces to determining how much message processing cost falls into each of these categories. While the BOSs might not have identical traffic patterns, interoperability dictates that each will communicate with the other with the same protocols and business logic. The traffic patterns are modeled by suitably choosing arrival rates and traffic routing probabilities, while business logic is modeled by choosing values for the visit ratios and mean service times.

For our system with only two railroads, the cross border inbound demand at BOS_A is triggered by messages generated at BOS_B and vice versa. We therefore have the following expressions for the hypothetical traffic intensities for different types of work at the two BOSs:

$$\rho_{a,CBO} = \beta_{a,b}\Lambda_a D_{a,CBO}$$

$$\rho_{a,INT} = (1-\beta_{a,b})\Lambda_a D_{a,INT}$$

$$\rho_{a,CBI} = \beta_{b,a}\Lambda_{\mathrm{b}} D_{a,CBI}, \qquad a, b\epsilon\{A,B\}, a \neq b.$$

Notice that the utilization due to the inbound cross border traffic at one BOS is driven by the message rate and the routing probability of the other BOS. Notice further that for values of $\beta_{a,b}$, approaching one, the intranetwork traffic is reduced. This is not surprising, since it implies that a larger fraction of the traffic is sent to the other BOS. We expect $\beta_{a,b}$ to be large if the BOS in question covers a small area with a lot of cross-border traffic, as might be the case for a regional railroad sharing borders with major railroads at a hub, as is the case in Chicago and environs.

The combined synthetic load on BOS_a is given by

$$U_a = \rho_{a,CBO} + \rho_{a,INT} + \rho_{a,CBI}, \; a\epsilon\{A,B\}.$$

Since the expression for the utilization contains a contribution due to traffic coming in from the other BOS, increasing incoming cross-border traffic will cause the average response time for all work done locally to increase.

Since we have assumed that the service discipline at each BOS is processor sharing (PS), we have the following expression for the total delay incurred in processing a message of type $k$ at BOS_a:

$$R_{a,k} = \frac{D_{a,k}}{1-U_a} \qquad a\epsilon\{A,B\}, k\epsilon\{CBI, CBO, INT\}$$

An authentication sequence goes through a sequence of steps involving local cross border processing and remote cross border processing. Other types of messages go through intranetwork processing. The number of visits to each BOS for processing is accounted for in the demand, which equals the visit ratio multiplied by the service time. The impact of authentication schemes and message traffic patterns on each of the delays will be reflected in the way $U_a$ and $D_{a,k}$ are computed. The average total processing time, excluding transmission delay, depends on the action being taken. For intranetwork communications, the message handling delay at BOS_a is given by

$$R_{a,INT} = \frac{D_{a,INT}}{1 - U_a}$$

For cross-border authentication, there is delay component due to local processing and a delay component due to remote processing. Hence, cross-border authentication delay is given by

$$R_{a,XBORDER} = \frac{D_{a,CBO}}{1 - U_a} + \frac{D_{b,CBI}}{1 - U_b} + T_N$$

where $T_N$ is a networking delay to be modeled separately.

We now consider the case where the driver of a train in network A is employed by network B, and suppose further that the train is traveling from network A to network B. The foreign driver must be authenticated by his or her home network. Authentication of the driver takes place while the train is in network A. Notice that there might be more than one train in network A whose drivers are employed by network B, though the reverse might not be true. For instance, more than one NJ Transit train may be running on the Amtrak corridor between New York and Trenton, but no Amtrak train will run on the North Jersey Coast Line which branches off the Northeast Corridor line. For this case, we assume that there is no change to the message traffic associated with a train crossing a network boundary, but that additional intranetwork processing is required. We suppose that the need for internetwork authentication of drivers employed by $b$ running on network $a$ increases the intranetwork message rate by some scale factor $1 + \varphi_{a,b}$ with $\varphi_{a,b} \geq 0$.Therefore,

$$\rho_{a,CBO} = \beta_{a,b}\Lambda_a D_{a,CBO}$$

$$\rho_{a,CBI} = \beta_{b,a}\Lambda_{\mathrm{b}} D_{a,CBI}, \qquad a, b \epsilon \{A, B\}, a \neq b$$

as before, while the utilization due to messages attributable to intra-network train traffic becomes

$$\rho_{a,INT} = (1 - \beta_{a,b})\Lambda_a D_{a,INT}(1 + \varphi_{a,b}) + (1 - \beta_{b,a})\Lambda_b D_{a,INT}\varphi_{b,a} .$$

We can rewrite this equation as:

$$\rho_{a,INT} = [(1 - \beta_{a,b})\Lambda_a(1 + \varphi_{a,b}) + (1 - \beta_{b,a})\Lambda_b\varphi_{b,a} ]D_{a,INT}$$

.

The first term accounts for the message flow due to intranetwork traffic, enlarged by overhead due to the need to initiate authentication at the other BOS. The second term accounts for the need for railroad $b$ to authenticate drivers employed by railroad $a$. If $a$=AMTRAK and $b$=NJC (North Jersey Coast Line), we have $\phi_{AMTRAK,NJC} > 0$ and $\phi_{NJC,AMTRAK} = 0$ , since (we assume) there are no Amtrak drivers on the North Jersey Coast

Line. The cross-border probabilities $\beta_{a,b}$ and $\beta_{b,a}$ are unaffected, because they only depend on the volume of trains crossing network boundaries, not on the affiliations of the drivers. From the foregoing, it follows that the total delays are of the same form as before, but with the utilizations changed to reflect the cost associated with authenticating "foreign" drivers.

## 5.5. Numerical Examples: BOS Utilizations and Delays

In this section, we illustrate the effect of varying $\beta_{A,B}$ from 0.1 to 0.5 on the utilizations and message processing delays for a given set of message rates per train. The utilizations at BOS_A fall linearly with $\beta_{A,B}$, while those at BOS_B increase linearly with $\beta_{A,B}$, as one might expect. One can also see how the authentication response times are affected by the shift of load from one BOS to the other, bearing in mind that authentication entails delays at both back office servers. The utilizations are shown in Figure 9 and the corresponding delays in Figure 10. None of the delays rises very steeply in over the range shown, because the traffic intensities and hence the (synthetic) utilizations are all below 70%.

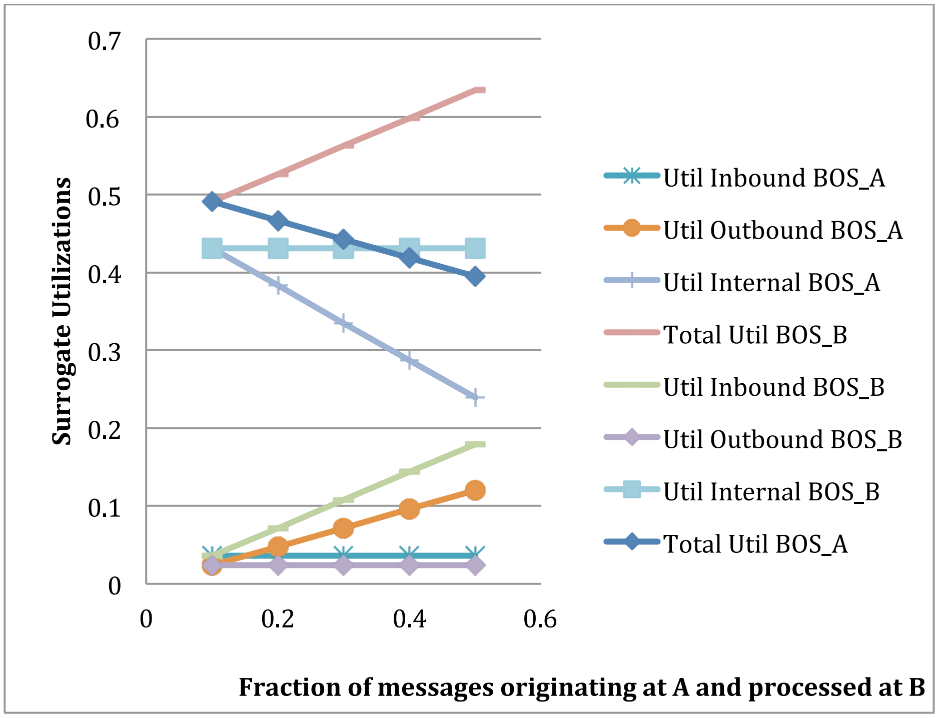

**Figure 9. Utilizations as a function of the fraction of OBU-BOS and BOS-OBU traffic related to cross-border authentication.**

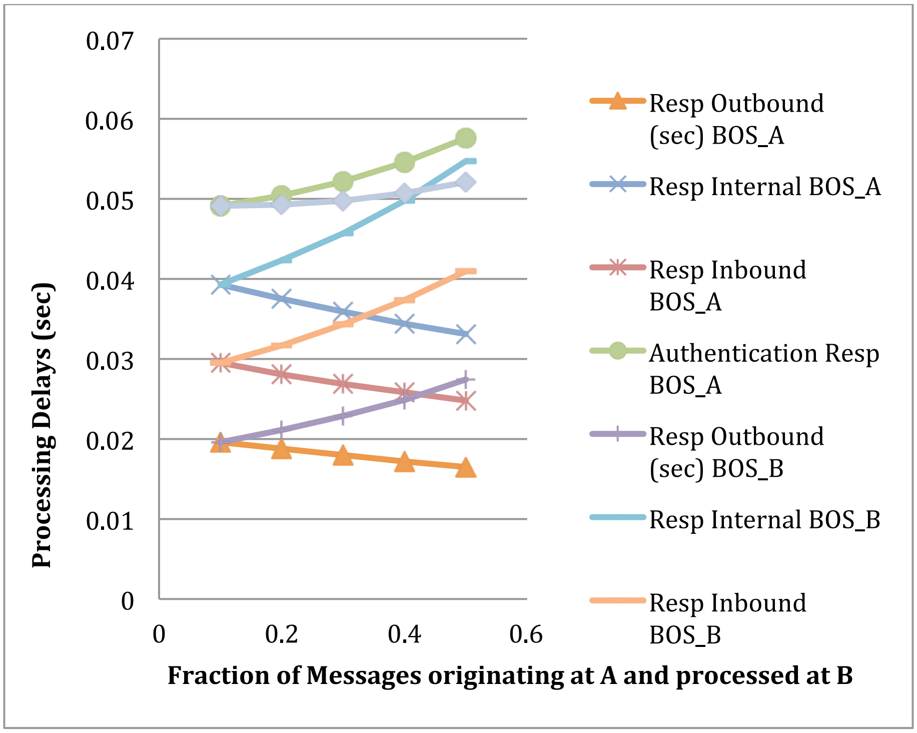

**Figure 10. Processing delays for OBU-BOS and BOS-OBU message traffic and cross-border authentication traffic.**

In the next two figures, we examine the effect of varying the number of messages per train per hour handled at BOS_A, while the corresponding number at BOS_B is held constant. All other parameters are the same as in the previous two figures, with $\beta_{A,B} = \beta_{B,A} = 0.1$. **Error! Reference source not found.** shows that the utilizations at BOS A and BOS B vary linearly with the number of messages per train per hour, with the latter varying less markedly than the former. **Error! Reference source not found.** shows that the response times vary less markedly at BOS_B than at BOS _, but are nonetheless increasing all round.

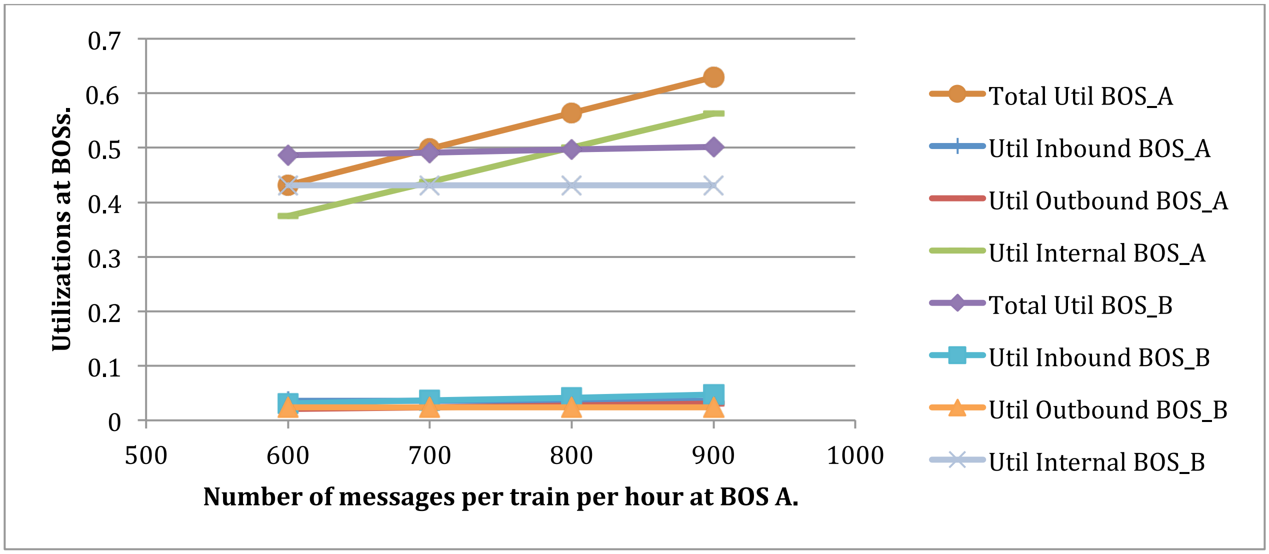

**Figure 11. BOS utilizations as a function of the number of messages per train at BOS A, with all other parameters held constant.**

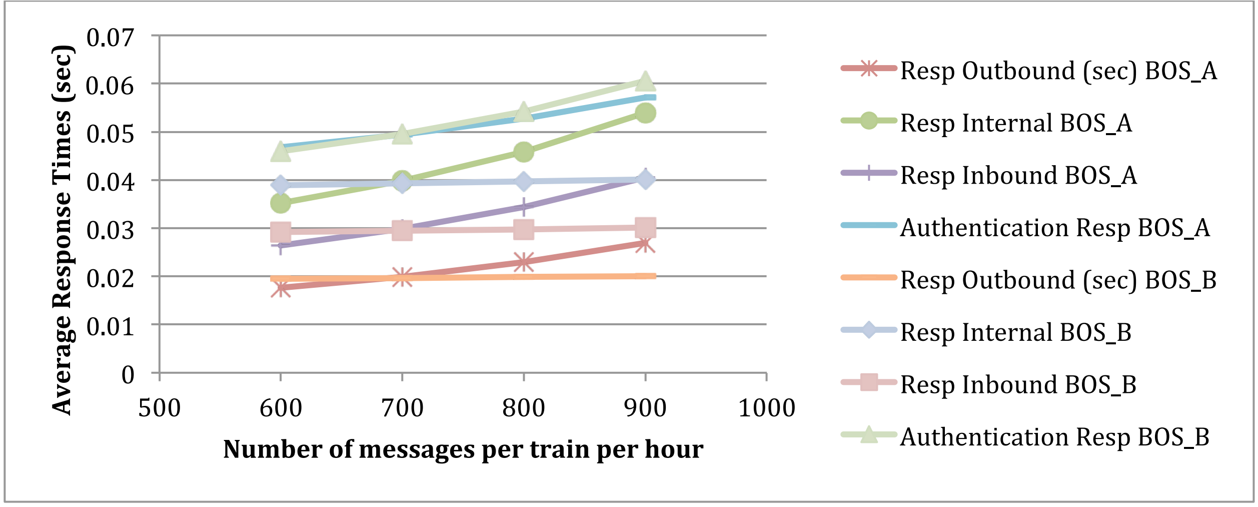

**Figure 12. Response Times as a function of the number of messages per train per hour at BOS_A, with all other parameters held constant.**

## 5.6. Ongoing Message Transmission – Status Updates

A locomotive's owner and its crew's employer may wish to be aware of the locomotive's current status. This would be forwarded by the host BOS to the BOS(s) belonging to the carrier that owns the locomotive and the carrier that employs the crew. These may all be the same, or there may be other combinations. These communications are not part of the message set specified in [S9352A], though they are derived from the contents of those messages. Therefore, there is no point in sending updates from BOS to BOS more frequently than the source BOS is updated. Updates from BOS to BOS could be less frequent, depending on operational considerations. This could be the subject of future work.

## 5.7. Discussion

Our goals in this work have been to characterize radio bandwidth usage and processing demands attributable to message traffic between trains' On-Board Units and Back Office Severs, and to model the delays associated with an authentication mechanism involving handshaking between trains and back office servers as trains approach railroad boundaries, even though draft standards as presently formulated do not call for such a mechanism. The amount of cross-border traffic and the consequent need for authentication vary with the distance between boundaries and how often trains cross them.

# Section 6. Conclusions

This report is our final delivery for the Grant Number FR-TEC-0006-11-01-00/20.321. This particular document is the first part of the document that describes the way used to divide the reporting of our work in investigating the identity management of proposed PTC systems. Section 2 Sections 4 and 5 described our performance studies.

Our investigation of the draft standard for communication between an OBU and its host BOS [S9352A] shows that the demand for bandwidth due to a train en route could be dominated by train position indication messages if the local BOS requests that they be sent often. By comparison, it appears that most other types of messages, whether they are sent from the OBU to the BOS or vice versa, occur somewhat infrequently on a train-by-train basis. Incidentally, poor training of engineers will result in more message traffic, since an OBU will notify the local back office server every time there is an operating violation of any kind, whether it is passing a signal or exceeding the local speed limit.

# Acknowledgements

This project has benefited from discussions with Volker Knollmann and Michael Smith at Siemens IMO in New York and Daniel Paulish at Siemens Corporation, Corporate Technologies (formerly Siemens Corporate Research) in Princeton, NJ.

# References

[Alstom2011] “Amtrak Introduces Alstom PTC Wayside Interface Unit along Northeast Corridor”,http://www.alstom.com/us/products-and services/transport/usptcmilestones/amtrakWIU2011/.

[Barker2009] E. Barker, W. Burr, A. Jones, T. Polk, S. Rose, and Q. Dang, *Recommendation for Key Management Part 3: Application-Specific Key Management Guidance,* NIST Special Publication 800-57, 2009.

[BCMP1975] Baskett, F. K. M. Chandy, R. R. Muntz, and F. Palacios. Open, closed, and mixed Networks of queues with different classes of customers. JACM 22(2), 248-260, 1975.

[Bondi2012] Bondi, A. B. Personal observation at the AMTRAK Wilmington control center.

[Cohn1958] Cohn, P. M., *Linear Equations,* Routledge & Kegan Paul, 1958.

[Cooper1981] Cooper, R. B., *Introduction to Queueing Theory, Second Edition,* North Holland, 1981.

[DahlBjork1974] Dahlquist, G., and Ake Bjork: *Numerical Methods,* Prentice Hall, 1974.

[Damodaran2006] Meledath Damodaran, Secure software development using use cases and misuse cases, Issues in Information Systems , Volume VII, No. 1, 2006.

[Duarte2008] F. Duarte, W. Hasling, W. Sherman, D. Paulish, R. Leao, E. Silve, and V. Cortellessa, Extending Model Transformations in the Performance Domain with a Node Modeling Library. Proc. WOSP 2008, 157-164, 2008.

[FCCapplicationHint2012] http://www.pacificcrest.com/library/AppNote_Applying_25kHz_FCC_License.pdf

[FRA2012] Federal Railroad Administration Report to Congress: Positive Train Control Implementation Status, Issues, and Impacts, August 2012. http://www.fra.dot.gov/downloads/08.2012_FRA_Report_to%20Congress_on_Positive_Train_Control.pdf

[Franks2009] G. Franks, T. Al Omari, M. Woodside, O. Das, and S. Derisavi, *Enhanced Modeling and Solution of Layered Queueing Networks,* IEEE Trans. SE 35(2), 148161, 2009.

[Groepler20xx] Gropler, O.: Braking Curves and Models for ETCS. Deutsche Bahn AG. *Standards submission.*

[GrossHarris1985] Gross, D., and C. Harris, *Fundamentals of Queueing Theory, 2nd Edition,* Wiley, 1985.

[Haller1995] N. Haller, *The S/Key One-Time Password System*, Network Working Group, Request for Comments: 1760, February 1995

[Hartong2006 ] Mark Hartong, Rajni Goel and Duminda Wijesekera,Use-Misuse case driven analysis of Positive Train Control, Proceedings of IFIP Int. Conf. Digital Forensics 15, 2006.

[Hartong2008] Mark Hartong, Rajni Goel and Duminda Wijesekera,Trust based secure positive train control interoperation, J. Transp. Secur. (2008) 1:211–228 DOI 10.1007/s12198-008-0019-7.

[Hartong2009] Mark Hartong, Secure Communication Based Train Control (CBTC) Operations, George Mason University, Ph.D. Thesis, Spring Semester 2009.

[Hunt2011] Hunt, Mat: *GATC Braking Curve Model.* http://hyperkahler.co.uk/gatc-braking-model.

[ITC-S9001_2011] Interoperable Train Control (ITC)-System Reference Architecture- S9001 version 1.3, 2011

[Kaufman2002] C. Kaufman, R. Perlman, M. Speciner, *Network Security: Private Communication in a Public World, Second Edition,* Prentice Hall, 2002.

[Ku2003] W. C. Ku, H. C. Tsai, and S. M. Chen, "Two simple attacks on Lin-Shen-Hwang's strong-password authentication protocol," ACM Operating System Re-view, vol. 37, no. 4, pp. 26-31, Oct 2003.

[Lamport1981] L. Lamport, Password authentication with insecure communi-cation, Communications of ACM 24 (1981) 28 – 30

[Lin2001] C. L. Lin, H. M. Sun, and T. Hwang, "Attacks and solutions on strong-password authentication," IEICE Transactions on Communications, vol. E84-B, no. 9, pp. 2622-2627, Sep 2001.

[MALD2004] Menasce, D. A.; Almeida, V. A. F.; and L. W. Dowdy. *Performance by Design.* Prentice Hall, 2004.

[Murphy2012] Murphy, C. The Internet of Things: Union Pacific shows the enormous potential of the instrumented, interconnected, analytics-sensitive enterprise. *InformationWeek* , pp. 16-21, August 13 2012.

[Pachl2002] Pachl, J., *Railway Operations and Control.* VTD Rail Publishing, 2002.

[Perrig2001]Adrian Perrig, Robert Szewczyk, Victor Wen, David Culler, J.D. Tygar, SPINS: security protocols for sensor networks, Proceedings of ACM MobiCom'01, Rome, Italy, 2001, pp. 189–199

[Perrig2002] Adrian Perrig, J.D. Tygar, Secure Broadcast Communication in Wired and Wireless *Networks*, Kluwer Academic Publishers Group, 2003.

[Perrig2005] Adrian Perrig, Ran Canetti, J. D. Tygar, and Dawn Song, "The TESLA Broadcast Authentication Protocol" (2005). Department of Engineering and Public Policy . Paper 62. http://repository .cmu.edu/epp/62

[S9001] S9001 ITC – System Reference Architecture, Version 1.4, Interoperable Train Control Architecture Team, August 2011.

[S9202] AAR Manual of Standards and Recommended Practices. Interoperable Train Control: Wayside Interface Interface Unit Requirements, Standard S-9202. Adopted 2012.

[S9352A] PTC Office-Locomotive Segment ICD, Release 2.10, 2/29/2012.

[S9352B] S-9352B Interoperable Train Control (ITC): Wayside-Locomotive Interface Control Document. DRAFT FOR COMMENT – DO NOT USE. Version DRAFT 3, 21 December 2010.

[S9352C] AAR Manual of Standards and Recommended Practices: Office Architecture and Railroad Electronics Messaging. S-9352C: ITC TIME AND LOCATION—INTERFACE CONTROL DOCUMENT (ICD). Adopted: 2012.

[Sandirigama2000] M. Sandirigama, A. Shimizu, and M. Noda, "Simple and secure password authentication protocol (SAS)," IEICE Transactions on Communications, vol. E83-B, no. 6, pp. 1363-1365, Jun 2000.

[Shimizu1990] A. Shimizu, "A dynamic password authentication method by one-way function," IEICE Transactions, vol. E73-DI, no. 7, pp. 630-636, Jul 1990.

[Shimizu1998] A. Shimizu, T. Horioka, and H. Inagaki, "A password authentication method for contents communication on the internet," IEICE Transactions on Communications, vol. E81-B, no. 8, pp. 1666-1673, Aug 1998.

[TseBross2009] Tse, T., and Brosseau, J., *Development of an Adaptive Enforcement Braking Algorithm for PTC Systems.* FRA Presentation to the TRB, December 2009. File *TRB_DEC09* available at *http://www.fra.dot.gov*.

[V-ETMS2010] Vital Electronic Train Management System(I-ETMS)-Concept of Operations, 24 March 2010,version 1.0

[Vincze2006] B. Vincze and G. Tarnai: Development and Analysis of Train Brake Curve Calculation Methods with Complex Simulation.

[Wang2005] Xiaoyun Wang, Yiqun Yin, and Hongbo Yu, Finding collisions in the full sha-1, In Proceedings of Crypto, August 2005

[Wei2010] Wei, ShangGuan and Cai Bai-gen, Wang Jing-jing and Wang Jian. Research and Analysis of ETCS Controlling Curves Model, *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICA CTE)*, v2-178-v2-181, 2010.

[Whittle2008] Jon Whittle, Duminda Wijesekera, Mark Hartong, "Executable misuse cases for modeling security concerns," icse, pp.121-130, Proceedings of the 30th International Conference on Software Engineering (ICSE '08), 2008

[WIU-AAR9202_2010] Interoperable Train Control Wayside Interface Unit Requirements Railway Electronics-AAR S-9202_2010

[WIU-AAR9202_2010] Interoperable Train Control Wayside Interface Unit Requirements Railway Electronics-AAR S-9202_2010

[Zhu2006] L. Zhu and B. Tung, *Public Key Cryptography for Initial Authentication in Kerberos,* Network Working Group, Request for Comments: 4556, 2006.